Adv Biochem Eng Biotechnol
DOI 10.1007/10_2014_301


# Cytochromes P450 for Terpene Functionalisation and Metabolic Engineering

**Irini Pateraki, Allison Maree Heskes and Björn Hamberger**

**Abstract** Plants have evolved the capacity to produce a striking array of specialised metabolites. Terpenoids are the oldest and most diverse class of such compounds and have attracted interest for industrial and pharmaceutical applications. The development of biotechnological alternatives for their production is the focus of intense research. Photosynthetic systems provide new strategies for autotrophic metabolic engineering. Focusing on cytochromes P450, involved in the functionalisation of the core terpene molecules, this review highlights the latest approaches in this field and looks towards recent discoveries that have the potential to shape the future of terpenoid bioengineering.

**Keywords** Cytochromes P450 · Biotechnology · Terpenoids · Production hosts · Pathway engineering · Terpenoid biosynthesis

**Abbreviations**

| | |
|---|---|
| P450 | Cytochrome P450 dependent mono-oxygenase |
| POR | NADPH dependent cytochrome P450 oxidoreductase |
| CYP71A1 to CYP71 clan | Example of classification of P450s into subfamilies to clans |

I. Pateraki · A.M. Heskes
Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Copenhagen, Denmark

I. Pateraki · A.M. Heskes · B. Hamberger
Center for Synthetic Biology bioSYNergy, Copenhagen, Denmark

I. Pateraki · A.M. Heskes · B. Hamberger
Copenhagen Plant Sciences Centre, University of Copenhagen, Copenhagen, Denmark

B. Hamberger (✉)
Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, Frederiksberg C, 1871 Copenhagen, Denmark
e-mail: bjoernh@plen.ku.dk

| | |
|---|---|
| EST | Expressed sequence tag |
| $Fsp^3$ | Fraction of carbon atoms in compound with $sp^3$ hybridisation |
| ER | Endoplasmatic reticulum |

## Contents



## 1 Introduction

Plant-derived natural products have been utilised by humans to improve health and nutrition throughout history. Currently, two thirds of the global population still rely on traditional medicines consisting mostly of plant crude extracts. At the same time, the annual spending on pharmaceuticals by the established market economies (representing one fifth of the world's population) is projected to grow by US $130 billion over the next five years (e.g., [116]). Interestingly, it seems that these industrialised, synthetic pharmaceuticals still have their origins in the natural world. Of the over 900 small-molecule drugs approved between 1981 and 2010, approximately two thirds mimic or are analogues of natural compounds, carry a natural pharmacophore or are naturally inspired or derived [88].

Terpenoids represent prime examples of natural bioactive molecules. The earliest evidence of use and consumption of plant species rich in terpenoids dates back 50,000 years to the Palaeolithic era, with two independent lines of research indicating that Neanderthal hominids self-medicated with what are today recognised medicinal plants (i.e., yarrow, chamomile; [46, 67]). Terpenoids fulfil two key criteria that are associated with compounds that have good potential as pharmaceuticals [73]. First, the $C_{15}$, $C_{20}$, and $C_{30}$ scaffolds of sesqui-, di-, and triterpenes carry an exceptionally high fraction of carbons in tetrahedral $sp^3$—versus planar $sp^2$—and linear

sp-hybridisation ($Fsp^3$). Second, they contain a high number of chiral centres (Fig. 1). In assays for inhibition of enzymes of the human liver, used as a proxy for toxicity of drugs due to off-target effects, the presence of these structural features is positively correlated with target specificity and low toxicity [72]. Yet, when compared with compounds of the alkaloid class, there is a discrepancy between the number of known plant terpenoids and those recognised in one of the 247 categories defined as medically active compounds in the Dictionary of Natural Products (Table 1; DNP23.1, accessed

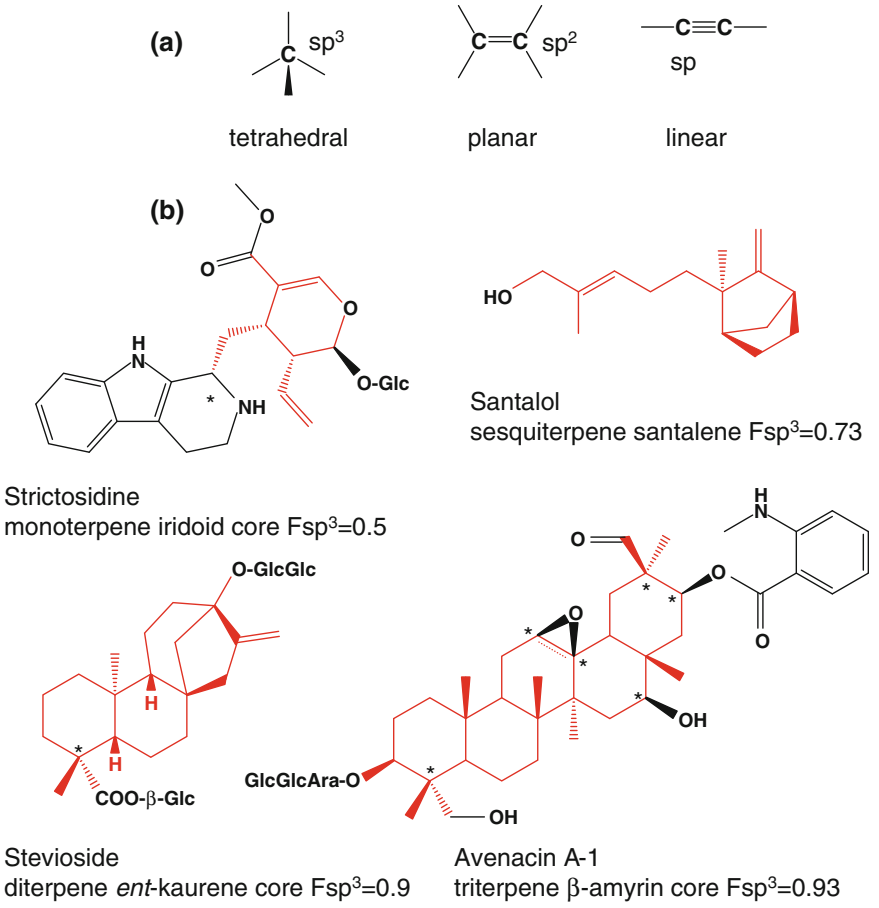


**Fig. 1 The impact of carbon hybridisation on potential structural diversity**. **a** Hybridisation of carbon orbitals and resulting geometry. *Note* Only $sp^3$ hybridised carbon atoms allow stereoisomers. **b** Selected examples of structurally complex terpenoids with increasing fraction of $sp^3$ hybridised carbon atoms ($Fsp^3$). In *red* Iridoid core, two stereocenters; santalene sesquiterpene, three stereocenters; *ent*-kaurene diterpene, six stereocenters; beta-amyrin triterpene, eight stereocenters; an *asterisk* indicates stereocenters in the structures formed during functionalisation. *Dashed line* indicates a double bond in the beta-amyrin backbone, replaced by an epoxide in avenacin. Examples given are: secologanin/monoterpene iridoid core in the terpenoalkaloid seco-iridoid strictosidine [77]; (E)-epi-β-santalol [29]; Steviol in the terpene glycoside Stevioside [91]; 12,13β-epoxy-16β-hydroxy-β-amyrin in the compound antifungal Avenacin A-1 [38]

**Table 1** Numbers and types of terpenoids currently deposited in the dictionary of natural products

| Type of terpenoid[a] | Compounds (DNP 23.1) | Classified as drug | Fraction with 2+ oxygen[b] (%) |
|---|---|---|---|
| Monoterpenoid | 4,129 | 28 | 92.2 |
| Sesquiterpenoid | 13,981 | 97 | 89.5 |
| Diterpenoid | 12,505 | 52 | 95.7 |
| Triterpenoid | 23,051 | 155 | 95.1 |
| Tetraterpenoid | 827 | 5 | 90.1 |
| Total | 54,493 | 337 | 93.5 |

[a] Including terpene-derived products within each class, such as apo- and nor-terpenoids
[b] Up to two oxygen atoms can be the result of water access and quenching of the carbocation during formation of the terpene backbone by terpene synthases

July 2014). Among the plant-based drugs approved, or with clinical trials launched in the period from 2000 to 2010, terpenoid-related compounds are dominated by only three scaffolds: Taxol (or paclitaxel), artemisinin, and vinblastine [14, 88, 103]. The key for the successful development of these scaffolds in clinical applications and establishment as treatment has been securing supply of the compounds through intense efforts in biotechnology or semi- or full chemical synthesis [23, 93, 128].

In addition to their potential as pharmaceuticals, bioactive terpenoids have attracted substantial interest as neutraceuticals, flavours and fragrances, and for their biopesticidal and antiherbivore activities (Fig. 1). This has led bioactive terpenoids to be considered as high-value natural products, setting this class of compounds apart from terpenoids that are used as fuels or feedstock for the chemical industry.

Regardless of the area of application, the availability of terpenoids for industrial purposes is, in general, hampered by laborious extraction and purification from their plant sources or limited economic competitiveness of chemical synthesis. Recently, biotechnological production has emerged as an alternative source of terpenoids for industrial applications. This technology involves the reconstruction of terpenoid biosynthetic pathways in host organisms that are suitable for large-scale cultivation of cell factories and holds the promise of being a biosustainable alternative to conventional production methods. The major requirement for successful engineering of the metabolic routes to complex target terpenoids in production hosts is knowledge of the enzymes involved in their biosynthesis.

Terpenoid biosynthesis involves two main classes of enzymes. The formation of the complex, often multicyclic scaffolds is catalysed by enzymes of the terpene synthase families [21]. The resulting scaffolds are then oxidatively functionalised through the action of cytochromes P450 (P450s). These enzymes carry out stereospecific hydroxylations, regularly contributing novel chiralities to the molecule, and which serve as molecular handles for further modifications, that is, linkage of sugar residues, alkylations, or esterifications. Further oxidation of hydroxyl groups to carbonyl and carboxyl groups is also a typical function of P450s. It should also be noted that introduction of up to two oxygen-containing functional groups into the terpenoid scaffold can be the result of water-quenching of the carbocation in the active site during formation by terpene synthases (e.g., [16, 131]).

This review highlights recent efforts and changes in the strategies for the discovery of plant enzymes with a focus on the extremely diverse family of P450s involved in oxidative functionalisation of high-value terpenoids. It also discusses advances in heterologous expression and metabolic engineering of selected pathways in the biotechnological hosts, *E. coli* and yeast, and also some key plant species. Finally, a brief future perspective is provided in which new discoveries which may shape the emerging field are discussed.

## 2 Genetic Diversity of Plant P450s as Driver of Terpenoid Diversity

Cytochromes P450 constitute the largest and functionally most versatile enzyme superfamily found in nature [48]. They are heme-containing enzymes, found in all kingdoms of life, noted for their extremely diverse amino acid sequences. They participate in many different biological processes, ranging from general biosynthetic functions (e.g., steroid biosynthesis) to specialised metabolism (plant defence molecules) and detoxification of xenobiotic chemicals such as drugs and pollutants [86]. The primary reaction they catalyse is the introduction of atmospheric oxygen into nonactivated carbon–hydrogen bonds [62]. For their catalytic activity, P450s require reducing equivalents, which are usually provided by NADPH or NADH, depending on the specific P450 enzyme. Inasmuch as P450s cannot receive electrons directly from these cofactors, different redox partners are engaged, such as membrane-bound FMN/FAD containing NADPH-dependent P450 oxidoreductases (POR) in the case of eukaryotic enzymes [62].

P450 enzymes are classified according to their sequence identity and not necessarily to their specific activity. Hence, the nomenclature reflects evolutionary relationships of the sequences [86]. Generally, P450s with a sequence identity higher than 40 % constitute members of one family, and sequence identity over 55 % defines members of a subfamily. In an effort to categorise the different P450 families in a higher order and make their nomenclature more comprehensive, the term ‘clan’ has been introduced [83, 84]. Clans are defined as groups of CYP families which segregate together on phylogenetic trees with members considered to have evolved from a common ancestor [86].

The diversity of P450 genes is vast, yet little is known about the evolutionary mechanisms driving the emergence of novel families. Due to their ubiquitous distribution, ancient P450s of terpenoid general metabolism are considered the genetic raw material for evolution. More than six decades after the first discovery of a P450, 127 families have been established in the plant kingdom, compared to 67 families in insects and 19 in vertebrates [85]. And within those families, over 100 subfamilies now carry functionally characterised P450s, of which 46 subfamilies, distributed over 8 of the 11 known plant clans, are involved in both general and specialised metabolism of terpenoids (see Appendix table for an update on details on plant species and references).

A characteristic of rapidly evolving families and subfamilies of P450s with roles in specialised metabolism is the occurrence of species- or genus-specific expansions called 'blooms' [33, 113]. This is particularly apparent in the CYP71 clan, where 20 out of 21 subfamilies are multimembered and have proliferated to such an extent that the previous definitions of family and subfamily have become problematic [85]. David Nelson coined the term 'tribe' for larger clades of subfamilies, in analogy to the rank of tribe between family and genus in taxonomy. This terminology had already been applied to the large group (i.e., tribe) of CYP71D found in the Euphorbiaceae, that consists of several phylogenetically distinct clusters of CYP71D and CYP726A subfamilies (Fig. 2; [132]).

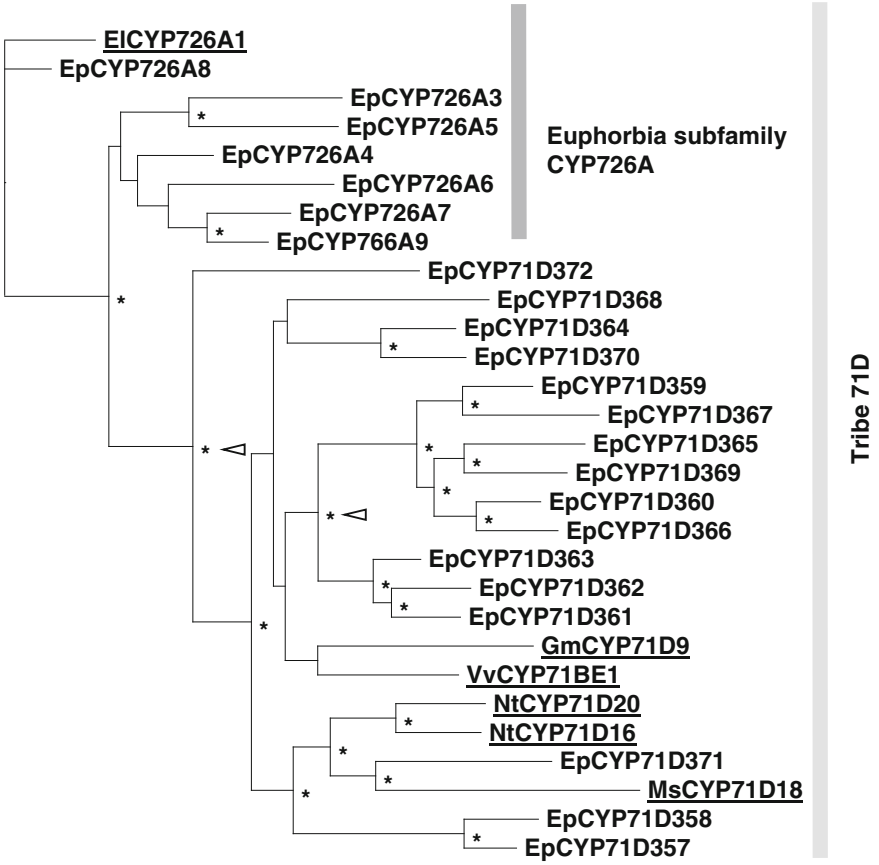


**Fig. 2** Species and plant family specific blooms, resulting in P450 multigene families (modified after: [132]; copyright American Society of Plant Biologists, www.plantphysiol.org, permission obtained). El, *Euphorbia lagascae*; Ep, *E. peplus*; Gm, *Glycine max*; Vv, *Vitis vinifera*; Nt, *Nicotiana tabaccum*; Ms, *Mentha spicata*. Asterisks indicate bootstrap values equal to, or above 80 %; triangles indicate phylogenetically distinct clades in *E. peplus*

The origin of such expansions of P450s is largely unknown, although it has been suggested that the fitness gain resulting from chemical diversification is one force driving the evolution of new P450s, with the original genes being recruited from the general metabolism [42]. Evidence of such a process can be observed in the rice genome, where five CYP701A homologues constitute a small multigene family, whereas only one, CYP701A6, is required for gibberellic acid biosynthesis. Its homologue CYP701A8, generated after duplication from general metabolism, has undergone neofunctionalisation and is now involved in the biosynthesis of rice antifungal phytoalexins [124]. A second example from rice is the grass-specific CYP51H subfamily, which contains nine members [87]. CYP51s are an ancestral P450 group typically found in eukaryotes in a single copy that catalyse a three-step demethylation of $C_{30}$ sterol constituents of membranes. Little is known about the function of these additional CYP51 members in rice, but one of the members from the subfamily in oat has been functionally characterised as the first oxidative enzyme in the biosynthetic pathway of avenacin saponins, glycosylated triterpene-derived defence compounds [38].

## 3 P450s Involved in Plant Mono- to Triterpenoid Metabolism

The discovery of an enzymatic route to a given compound is critical if its biotechnological production is to be achieved. The development and maturation of next-generation sequencing techniques over the last decade, pioneered by 454 sequencing [76], has significantly increased our ability to perform this task. Several genomics and functional genomics projects focused on medicinal plants accumulating bioactive terpenoids have released vast nonmodel plant sequencing resources to the public, including the Medicinal Plant Genomics Resource (http://medicinalplantgenomics.msu.edu/), the Canadian PhytoMetaSyn Initiative (http://www.phytometasyn.com/), and the 1KP Plant project (http://www.onekp.com/project.html). In the wake of this radically increased availability of sequence data, creative and efficient approaches are emerging to identify the fraction of genes relevant for terpenoid biosynthesis, particularly in regard to the identification of P450s as these exist in high numbers in plant genomes. The following selected illustrative approaches have succeeded in the identification of specific P450s by integrating these sequencing data with proteomic and metabolomic data, performing deep comparative transcriptomics (e.g., coexpression analyses) or by exploiting the fact that genes in terpenoid biosynthetic routes may occur physically clustered at the genomic level. Considering such clustering recently prompted a genomewide scan for physically linked pairs of terpene synthases and P450s in 17 monocots and eudicots. In several plant species colocalisation of terpene synthases was found, with significant bias towards specific families of P450s, providing a strong incentive for functional investigation [12]. Generally, following identification, enzymes are functionally characterised in heterologous systems, such as the microbes

*E. coli* and yeast, but also with the classical higher eukaryotic ovarian-derived insect *Spodoptera frugiperda* 9 (*Sf*9)-cell line/Baculovirus system and increasingly through transient expression in the plant *Nicotiana benthamiana* using *Agrobacterium* infiltrations.

## *3.1 Monoterpenoids*

Monoterpene indole alkaloids are a large group of plant-specialised metabolites with broad bioactivities ranging from anticancer and anti-inflammatory to antimicrobial properties. Biosynthesis of the central metabolite strictosidine involves oxidation of the initially volatile monoterpene geraniol to 8-hydroxygeraniol, cyclisation of the aldehyde 8-oxogeranial to an iridoid, and later formation of loganic acid, which is converted to loganin through an O-methyl transferase reaction. After oxidative ring-opening of loganin yielding secologanin, strictosidine is finally formed by coupling with tryptamine. The full biosynthetic pathway is suggested to include approximately 10 enzymes that catalyse the oxidation, reduction, glycosylation, and methylation reactions [77].

To identify missing steps in this complex biosynthetic pathway, deep transcriptome resources were developed for the key monoterpene-indole alkaloid-producing species, Madagascar periwinkle (*Catharanthus roseus*), the source of the anticancer monoterpene-indole alkaloid vinblastine (e.g., [39, 77]). Databases of annotated transcripts from various monoterpene-indole alkaloid-producing plant species were searched to identify candidate genes. Virus-induced gene silencing in plants of selected candidates and targeted metabolite analysis identified CYP72A224 by the accumulation of the presumed precursor of loganic acid, 7-deoxyloganic acid and a reduction of more complex monoterpene-indole alkaloids, indicating a function of this gene early in the pathway [104].

The last four unknown steps in the pathway to strictosidine were finally elucidated in 2014 using a gene coexpression analysis approach in which 23 transcriptomes were compared from different tissues, cell types, and treatments along with the proteomes of epidermal and mesophyll protoplasts isolated from *C. roseus* leaves [77]. By comparing different tissues, this approach was able to take into account cell-specific and spatially distant localisation of branches of the pathway. Biochemical characterisation of candidates resulted in identification of the enzymes catalysing formation of the iridoid-precursor, 8-oxogeranial, oxidation of the iridoid to 7-deoxyloganetic acid followed by glycosylation to loganic acid, and final oxidation to loganic acid (Fig. 3). Functional characterisation of the recombinant enzymes in *E. coli* established an 8-hydroxygeraniol dehydrogenase of the zinc-containing alcohol dehydrogenase family and a UDP-glucose-dependent glycosyltransferase involved in glycosylation of 7-deoxyloganetic acid to 7-deoxyloganic acid. Coexpression in yeast of CYP76A26 with the Arabidopsis POR1 established this CYP71 clan member as a multifunctional enzyme catalysing successive hydroxylation/dehydration steps en route to 7-deoxyloganetic acid. The mechanism

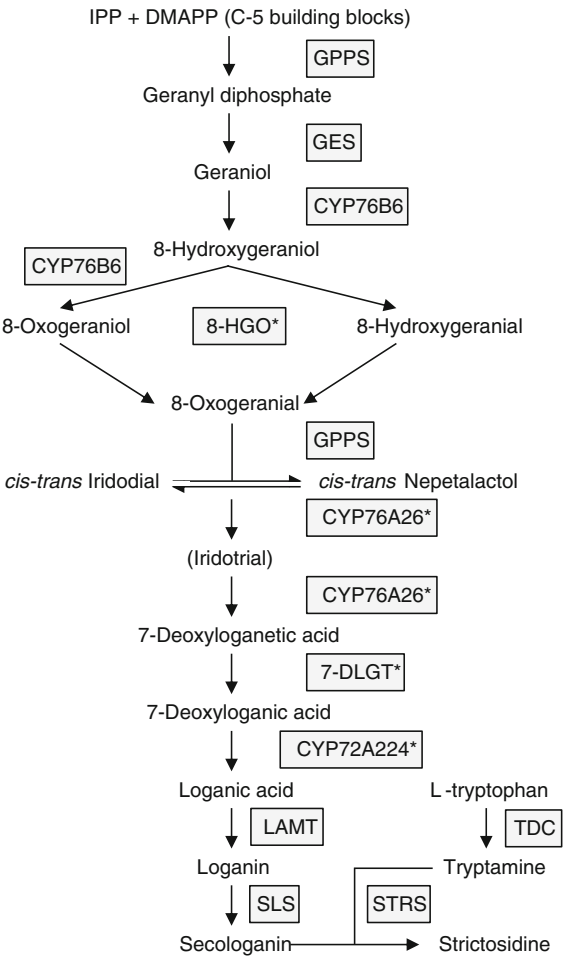


**Fig. 3** Heterologous expression of the strictosidine biosynthetic pathway in *N. benthamiana* (modified after: [77]). Enzymes specifically discussed in the text are marked by an *asterisk*. IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPPS, geranyl diphosphate synthase; GES, geraniol synthase; CYP76B6 (*C. roseus*), geraniol 8-oxidase; 8-HGO, 8-hydroxygeraniol oxidoreductase; IS, iridoid synthase; CYP76A26 (*C. roseus*) iridoid oxidase; 7-DLGT, 7-deoxyloganetic acid glucosyl transferase; CYP72A224 (*C. roseus*), 7-deoxyloganic acid hydroxylase; LAMT, loganic acid O-methyltransferase; SLS, secologanin synthase; STRS, strictosidine synthase; TDC, tryptophan decarboxylase

resembles the oxidation of *ent*-kaurene to *ent*-kaurenoic acid in gibberellin biosynthesis by members of the subfamily CYP701A of general metabolism, typically rooting the CYP71 clan.

Functional characterisation of the previously identified CYP72A224 [104], by expression in yeast and transient expression in *N. benthamiana* following *Agrobacterium* infiltration showed that this enzyme converted 7-deoxyloganic acid, but not the aglycone derivative of 7-deoxyloganic acid (7-deoxyloganetic acid) into loganic acid, and indicated that the glycosylation precedes the 7-hydroxylation. From the perspective of pathway discovery, this example of recruitment of the indole alkaloid pathway genes from three highly divergent families, zinc-containing alcohol dehydrogenases and P450s of the divergent clans CYP71 and CYP72, shows that enzymatic steps of core-pathways are not necessarily encoded by related

members of defined subfamilies. It remains to be tested whether other members of the relevant subfamilies can further increase chemical diversity of the individual steps. For functional characterisation of the enzymes involved, it also highlights the benefit of complementary host systems for heterologous expression. Finally, the availability of the enzymes catalysing all individual steps allowed the team to reconstitute the entire pathway to strictosidine transiently in the *N. benthamiana* system, demonstrating feasibility of biotechnological production [77].

## *3.2 Sesquiterpenoids*

(+)-Nootkatone is a volatile sesquiterpenoid found in traces in grapefruit epicarp (flavedo) and as the main component of Nootka cypress heartwood essential oil (*Cupressus nootkatensis*). It is considered to be a high-value compound due to its use as an aroma compound and its potential as an environmentally safe biopesticide and insect repellent [95]. Hence, the biosynthesis of nootkatone has attracted substantial attention over the last years, leading to several stepwise improvements in its biotechnological production. Nootkatone carries three stereocenters introduced by its bicyclic hydrocarbon scaffold, valencene, and a keto-function, indicating a potentially simple biosynthetic route. As a terpene synthase from 'Valencia' orange (*Citrus sinensis* cv. Valencia*)* had been functionally characterised over a decade ago as an efficient converter of farnesyl diphosphate to valencene [114], several groups embarked on the quest to discover enzymes for the oxidative conversion to nootkatone. Chicory is well known to accumulate a diverse spectrum of oxidised sesquiterpenes in its root. The finding that microsomal enzymes from chicory root could metabolise valencene led to the discovery of the CYP71 clan member CYP76AV8 in a root transcriptome, based on homology to terpenoid-oxidising enzymes from other species. Biochemical characterisation of the enzyme showed high substrate promiscuity and conversion of valencene, resulting in predominant accumulation of the alcohol nootkatol, with traces of nootkatone [17]. During further testing of enzymes of the CYP71 clan from species known to accumulate oxidised sesquiterpenoids, a P450 catalysing efficient oxidation of valencene towards nootkatone was finally discovered in a targeted deep transcriptome derived from Nootka cypress wood tissue, consistent with detection of nootkatone synthesis in this tissue. The founding member of a new P450 subfamily, CYP706M1, was shown to catalyse the formation of nootkatone (see the section on yeast engineering for details). Interestingly, the degree of amino acid identity of CYP706M1 compared to the previously identified P450s with valencene oxidase activity did not exceed 30 % [18], indicating that the evolution of specific sesquiterpene oxidases may occur through independent events. Furthermore, the limited success in replacement of the required specific two-step oxidation to nootkatone with P450s nonnative to this pathway highlights the advantages of dedicated species- and tissue-specific transcriptome libraries for P450 discovery.

The recent identification of a group of P450s responsible for the oxidation of sesquiterpenes in sandalwood provides another example of the utility of tissue-specific transcriptome mining for P450 gene discovery. Sandalwood (*Santalum album*) is a hemiparasitic tree species of commercial interest for the fragrance industry due to the accumulation of highly prized sesquiterpene-based sandalwood oil in its xylem tissue. Mining of the transcriptome of the xylem of *S. album* led to identification of ten P450s constituting two distinct clades of the CYP76F subfamily. Consistent with a function in biosynthesis of sandalwood oil, the genes were found to be highly coexpressed in *S. album* xylem tissue [29]. Functional characterisation in yeast of nine of the ten recombinant P450s showed a distinct but overlapping product profile for each clade, with some oxidised sesquiterpenes alcohols detected in both. The produced bergamotol and santalol sesquiterpenes were all monohydroxylated at the terminal carbon of the isoprenyl side-chain, and the overall profile was shown to match the composition of the sandalwood oil, yet with different proportions of the individual compounds [29].

## *3.3 Diterpenoids*

In general, the increasing molecular weight of terpenoids towards the $C_{20}$-based diterpenoids and the resulting loss of volatility of the core skeleton and changes in physicochemical properties leads to molecules with altered pharmaceutical, neutraceutical, and other industrially relevant properties [11]. One example of a group of high-value diterpenoids is the conifer diterpene resin acids, which are key components of high-end inks and tackifiers, used in formulating adhesives. To elucidate conifer diterpene resin acid biosynthesis, six divergent P450s from loblolly pine were selected based on remote similarity to families involved in triterpenoid phytohormones brassinosteroid metabolism (CYP90), or metabolism of Taxol (CYP725) of the CYP85 clan. Functional testing in yeast of all candidates allowed partial characterisation of CYP720B1 as a multisubstrate enzyme catalysing two consecutive oxidation steps leading from diterpene alcohols via the corresponding aldehyde to four structurally different resin acids [101]. Subsequent development and mining of large-scale transcriptome resources of nearly 1,000,000 pine and spruce ESTs revealed an extraordinary degree of multiplicity in the CYP720B subfamily specific for conifers. Recombinant Sitka spruce CYP720B4 was functionally characterised through complementary approaches in vitro, in vivo, and through RNA interference in planta and the enzyme was found to catalyse up to three oxidative steps of a panel of 24 diterpene olefins, alcohols, and aldehydes to varying degree [43].

To address the general lack of a standardised reference system for discovery and rapid functional annotation of genes involved in diterpenoid specialised metabolism in nonmodel plant systems, Zerbe and coworkers selected specific organs, tissues, or cell types for deep transcriptome sequencing, guided by targeted

metabolomics, [132]. Feasibility of the generalised strategy was demonstrated by mining transcriptomes of 10 plant species accumulating high-value diterpenes with custom-built libraries for diterpene synthases and P450s. Several hundred sequences for terpenoid-related P450s were initially discovered in close to 1,000,000 predicted unique transcript reads. Phylogenetic analyses of these genes for blooms resulting specifically in novel P450 multigene subfamilies restricted to target plant orders, families, or species, effectively reduced the number of candidate genes. A particularly striking example of such a bloom is the Euphorbiaceae specific expansion in the CYP71 family. Based on the expression of the 22 members of this group they were implicated in the biosynthesis of the complex macrocyclic diterpenes of the lathyrane- and ingenane-type characteristic of the *Euphorbia* species [132]. This hypothesis received independent support through analysis of genomic clustering of sequence homologues in the genome of another member of the spurge family, castor bean (*Ricinus communis*), and indications that the P450s are involved in oxidative decoration of simple macrocyclic diterpenoids in that species [58].

Within model plant species, the metabolism of complex diterpenoids is perhaps best understood in rice (*Oryza sativa*), where functional roles of structurally divergent families of diterpenoids as phytoalexins, phytoanticipants, and allelochemicals are well established in plant–microbe and plant–plant interactions [56, 97]. In contrast to nonmodel species where genomic resources are lacking, the keys to the rapid and comprehensive elucidation of diterpenoid pathways in rice have been the physical organisation of groups of biosynthetic genes in clusters and their coordinated expression in response to environmental stimuli or elicitation. For example, four expanded subfamilies of the CYP71 clan are found in clusters (chromosome 2, CYP76M, CYP71Z; chromosome 4 CYP99A, and chromosome 6, CYP701A) with two clusters also carrying diterpene synthases (reviewed in [107]). This distribution of the functional enzymes of rice diterpene metabolism over several subfamilies indicates independent evolutionary events leading to their emergence. It was suggested that the observed pattern arose early in the grass family of plants (Poaceae; [107]), consistent also with clustering of the founding member *S. bicolor* CYP99A1 [7] on chromosome 6, and a CYP99A-paralogue on chromosome 1 with sequences potentially encoding terpene synthases (http://phytozome.jgi.doe.gov, Phytozome10, Sorghum bicolor v2.1 genome).

## *3.4 Triterpenoids*

The oxidative functionalisation of triterpenoids involves evolutionary distant P450 enzymes found across four of the known eleven plant P450 clans: CYP51, CYP71, CYP72, and CYP85. In both oat (*Avena strigosa*) and liquorice (*Glycyrrhiza* spp.), discovery of duplications and divergence of genes originating in the metabolism of triterpene sterols (CYP51G) and diterpene gibberellins (CYP88A) indicated potential neofunctionalisation of newly emerged subfamilies. Indeed, liquorice CYP88D6 of the CYP85 clan was found to catalyse C-11 oxidation of β-amyrin,

representing the first committed step in the biosynthesis of the natural sweetener and pharmacologically active glycyrrhizin. It was initially discovered by narrowing down unique ESTs encoding candidate P450s by transcript profiling in the glycyrrhizin-accumulating underground organs (stolons and roots) and the aboveground glycyrrhizin-free organs (leaves and stems). The less obvious and evolutionary distant member of the CYP72 clan, CYP72A154, performing the subsequent oxidation of C-30, was later identified by expression profiling to find genes whose expression matched the organ-specific accumulation of glycyrrhizin and testing of identified candidates for enzymatic activity against 11-oxo-β-amyrin. Both CYP88D6 and CYP72A154 were characterised in vitro using microsomes from recombinant *Sf*9 insect cells, complemented by in vivo coexpression of the β-amyrin synthase, cytochrome P450 reductase (POR) with the P450s in yeast [111, 112].

Oat accumulates avenacins, glycosylated β-amyrin–derived triterpenoids conferring resistance to microbes. Even with limited biochemical knowledge, or evidence of the types of enzymes involved, linkage of biosynthetic genes had already been cautiously suggested 15 years ago based on genetic evidence of mutant lines, making this biosynthetic cluster an archetype for terpenoid biosynthesis [96]. The first gene involved in the oxidative decoration of avenacins, and encoding a P450, was indeed identified in a cluster with the triterpene β-amyrin synthase [98]. Both genes were shown to be coexpressed specifically in the epidermal cells of oat root tips, consistent with the localised accumulation of avenacin. Heterologous expression in yeast of CYP51H10 demonstrated catalytic activity converting β-amyrin to an oxidised triterpene. Yet, the conclusive structural elucidation of 12,13β-epoxy-16β-hydroxy-β-amyrin as intermediate of avenacin was only achieved after transient coexpression of the β-amyrin synthase with CYP51H10 in *N. benthamiana* following *Agrobacterium* infiltration and purification of the product for structure determination by NMR [38, 61].

Triterpene saponins from plants such as ginseng and liquorice have attracted interest due to their pharmacological properties. Substantial understanding of the biosynthetic routes of these compounds was gained over recent years from the model legume *Medicago truncatula*, a rich source of diverse saponins. A coexpression analysis based on methyl-jasmonate elicitation of triterpenoid biosynthesis, and correlation with expression of β-amyrin synthase was used to identify candidate P450s, which were reduced to only a few candidates by homology to P450s involved in terpenoid metabolism [82]. Indeed, *M. truncatula* CYP716A12, identified through homology with the conifer CYP720B, became the first functionally characterised member of this subfamily [35]. After initial demonstration of a functional enzyme in the microsomal fraction when expressed in the *Sf*9 system, heterologous expression of CYP716A12 in yeast together with a POR and three distinct triterpene synthases yielded oleanolic acid, ursolic acid, and betulinic acid from β-amyrin, α-amyrin, and lupeol, respectively. Similar activity of CYP716A homologues from grape towards triterpenes was also demonstrated, indicating a possible conserved functionality within the CYP716A subfamily. This groundbreaking work spurred a burst of discoveries of novel triterpene-oxidising members in the CYP716 family, such as in ginseng (*Panax* spp.), the African shrub *Maesa*

*lanceolata*, or the medicinal plant *Bupleurum falcatum* from the parsley family (Apiaceae; [44, 79, 80]) and also opened the door to combinatorial biosynthesis of triterpenes (see below in the section on perspectives).

## 4 Expression Systems: From Functional Testing of Recombinant P450s to Production Platforms

The biosynthesis of plant terpenoids by microbes for commercial production features several advantages over extraction from natural sources. These include: simple compound purification due to a limited chemical background, well-established genetic tools for metabolic engineering, and rapid generation of stable lines for scalable fermentation. However, the heterologous expression of functional terpene synthases and P450s, balancing of metabolic pathways, compartmentalisation of biochemical pathways, and metabolic channelling, are factors that currently need consideration and optimisation for every novel pathway and target compound where robust production lines and cost-effective biosynthesis are required. For example, plant mono- and diterpene synthases are plastidial enzymes in plants. Expression in *E. coli* or yeast hosts, which do not carry plastids, generally requires removal of the N-terminal plastid targeting peptide sequence of the native protein to yield a pseudomature form, as it could possibly interfere with proper protein folding and optimal expression [25, 106].

The expression of plant P450s in microbial hosts presents a specific set of challenges. Plant P450s are membrane-anchored enzymes localised at the endoplasmatic reticulum (ER) of the cell, making functional expression in prokaryotes problematic as these hosts do not contain this cellular compartment. Typical approaches to this problem involve removal or modifications of the N-terminal membrane anchor [8, 59, 66, 109]. Additionally, efficient electron transfer to recombinant P450s is required for function and this is normally achieved with the partner POR enzyme, colocalised and anchored at the ER [52]. A range of strategies for optimisations has been considered (summarised in [100]).

The following section provides examples of the successful expression of plant P450s in microbial hosts and a discussion of how arising problems have been overcome.

### *4.1 Bacterial Hosts*

An example illustrating some of the difficulties encountered during expression of P450s in bacterial hosts is the semisynthetic artemisinin project [93]. Artemisinin is a plant natural product of the sesquiterpene lactones class, produced in the trichomes of *Artemisia annua* [9]. Despite its designation as the most potent antimalarial drug and its effective use in antimalarial combination therapies (ACTs), the

supply was unstable due to the unreliable growth of *A. annua* crops. Thus, efforts have been undertaken over the last 10 years to develop an alternative biotechnologically produced source of artemisinin. The first microbial choice was *E. coli* where high titres (up to 27.4 g $L^{-1}$) of amorpha-4,11-diene (the sesquiterpene olefin precursor of artemisinin) were achieved by expression of the *A. annua* amorphadiene synthase (ADS) and a heterologous mevalonate pathway [121]. However, the biosynthetic pathway of artemisinin contains a single P450, CYP71AV1, which converts amorpha-4,11-diene to artemisinic acid or dihydroartemisinic acid through three consecutive oxidations [102, 118, 120]. Expression of the *A. annua* CYP71AV1 native enzyme in *E. coli* resulted in no detectable activity (in vivo or in vitro; [20]), creating a major obstacle for the successful use of *E. coli* as a production host.

To achieve activity of CYP71AV1 in *E. coli* multiple modifications and adjustments were required including: codon optimisation, manipulation of the N-terminal transmembrane domain, use of the native POR enzyme from *A. annua*, adjustment of the expression vectors, and also optimisation of the *E. coli* strain and culture conditions. Further efforts were undertaken to improve the yield, such as replacement of the anchor of the plant P450 with that of a heterologous CYP from an organism phylogenetically closer to bacteria (e.g., Candida tropicalis; [20]), or modifications of the anchor mimicking that of a well-expressing P450 [8]. Finally, with these changes, a significant amount of oxidation products could be detected (553 mg $L^{-1}$); however, only a fraction of these was artemisinic acid.

A different approach for the synthesis of artemisinic acid was followed by Dietrich et al. [30] with the hope of overcoming the low yields and the formation of by-products observed by Chang and coworkers [20]. Instead of the *A. annua* enzyme, a soluble substrate-promiscuous $P450_{BM3}$ from *Bacillus megaterium* (CYP102A1) was used to generate a new route towards artemisinic acid synthesis [30]. The wild-type $P450_{BM3}$ enzyme catalyses the hydroxylation of long-chain saturated fatty acids and has a very high turnover rate compared to other known P450s [90]. The wild-type enzyme was not active against amorphadiene and required several modifications for the enzyme to accept amorphadiene as a substrate. With the help of computer modelling, specific mutations were designed to increase the size of the active site binding pocket, to improve the enzyme specificity and for increasing the product titres (saturation mutagenesis). This semibiosynthetic strategy resulted in 250 mg $L^{-1}$ of artemisinic epoxide which can be converted through dihydroartemisinic acid to artemisinin [30]. The titres achieved by this method were higher than the previous ones using the native enzyme (CYP71AV1) expressed in the same host. This was the first time that the properties of a $P450_{BM3}$ were modified by directed mutagenesis and it seems that this strategy holds promise for different substrates and systems.

An important advantage of $P450_{BM3}$ is that it is composed of a fusion between a FAD–FMN-containing NADPH oxidoreductase and a functional P450, forming an entire monooxygenase system in one soluble 119-kDa polypeptide [81]. Indeed, the enzyme exhibited the highest catalytic activity ever detected in a P450, likely due to more efficient electron transfer [130]. Additionally, $P450_{BM3}$ has proven amenable

to targeted mutagenesis that can result in conformational changes and substrate specificity enabling novel P450 reactions for a diverse range of pathways and compounds [15, 22]. This case highlights the possibility of using existing enzymes for the catalysis of reactions leading to commercially interesting products whose pathways are not yet discovered or for the synthesis of new-to-nature products.

A further example where a bacterial enzyme has been used efficiently for the hydroxylation of a plant natural compound is in the pathway towards perillyl alcohol, an anticancer agent generated by the hydroxylation of limonene, a monoterpenoid found in several plants. Alonso-Gutierrez and coworkers [2] produced approximately 100 mg $L^{-1}$ perillyl alcohol from an *E. coli* strain engineered to produce high levels of the precursor limonene via a heterologous mevalonate pathway and the corresponding terpene synthase [2]. The hydroxylation of limonene was achieved by a well-characterised P450 isolated from Mycobacterium HXN-1500, coupled with a ferredoxin, and a ferredoxin reductase [122]. The specific P450 was identified after the screening of 1,800 bacterial strains for their ability to hydroxylate L-limonene at the 7 position [122].

Inspired by progress with the $P450_{BM3}$ enzyme, several artificial POR–P450 fusions are known today. One of these examples regards one of the most noteworthy success stories of the terpenoid world, Taxol. Taxol, a potent anticancer agent [34], is naturally found in the cork of the Pacific yew tree in very small amounts (more than 500,000 mature Pacific yew trees were felled for extraction of 130 kg of Taxol, required for preclinical studies; [69]), and today it is produced from plant cell cultures or by chemically converting deacetylbaccatin III (isolated from *Taxus baccata* needles) into Taxol. Although production today is more stable and efficient, the price of the drug still remains high [69]. Thus, several efforts have been implemented to approach biotechnological solutions for the synthesis of this drug. It is predicted that Taxol biosynthesis requires 19 distinct enzymatic steps from GGPP [24], with approximately half of these reactions being catalysed by P450s [55]. A significant number of the enzymes participating in the pathway have been characterised, however, there still remain several enzymatic steps missing including some potentially catalysed by P450s. Although the majority of known P450s involved in Taxol biosynthesis have been expressed in yeast or insect cells for functional characterisation [49–51, 108], there are relatively few reports regarding expression of Taxol-related P450s in microorganisms for production purposes [53, 68].

The work of Ajikumar and coworkers [1] describes the optimisation of the MEP pathway in *E. coli* for the synthesis of high titres of taxadiene (1 g $L^{-1}$ was achieved in fed-batch bioreactor fermentation) and the expression of the first P450 of the pathway, taxadien-5a-hydroxylase, which oxidises taxadiene to taxadien-5a-ol [51]. A codon-optimised taxadien-5a-hydroxylase was expressed, after N-terminal transmembrane domain engineering, as a fusion with the *Taxus* POR enzyme [1]. A number of fused enzymes (taxadien-5a-hydroxylase with *Taxus* POR) had been generated using different versions of the P450 transmembrane domain. One of the generated chimeric enzymes was able to convert 98 % of taxadiene to taxadien-5a-ol at a yield of 58 mg $L^{-1}$ and equal amounts of the by-product 5(12)-Oxa-3(11)-cyclotaxane. Although taxadien-5a-hydroxylase was suggested to be a bottleneck of the pathway

[28], the applied strategy made it possible to express the recombinant enzyme in *E. coli* functionally and to produce relatively high titres of its product. Despite the complications encountered regarding P450 expression, the observed titres were approximately 2,400-fold higher than the ones observed previously from yeast [28]; for further details refer to the relevant chapter in this volume). As all Taxol-related P450s have high sequence similarity, a refined approach inspired by this progress could potentially lead to functional expression of the next biosynthetic steps in *E. coli* [55].

CYP97 is a family of plant P450s that has been successfully expressed in microbial hosts. Enzymes from this family are involved in the hydroxylation of carotenoids for the synthesis of lutein (from α-carotene) or zeaxanthin (from β-carotene; [99, 119]), and are some of the few P450s natively localised to the plastids. In Arabidopsis, eleven P450s have been predicted/identified as plastidial targeted including CYP97A3 (carotene β-ring hydroxylase), CYP97B3 (carotene β-ring hydroxylase), and CYP97C1 (carotene ε-ring hydroxylase; [57, 99, 110]). Homologous enzymes have also been found in rice, soy bean, and pea [119]. As lutein is an important ingredient (as antioxidant and natural colorant) in a number of pharmaceuticals, food supplements, or cosmetics, efforts have been made to produce it biotechnologically.

The group of Eleanore Wurtzel (City University of New York, USA) succeeded in producing 500 μg of lutein per gram of pelleted cells by overexpression of the CYP97A4 and CYP97C2 enzymes from rice in *E. coli* cells harbouring the pathways for α- and β-carotene [127]. In plant plastids, the nonheme, diiron enzyme, HYD4, performs exactly the same carotene β-ring hydroxylation reaction as the CYP97A. To test the redundancy of these two enzymes, attempted replacement of CYP97A with the HYD4 in planta, but also in the bacterial cells, failed in lutein production, which was only possible by the coexpression of CYP97A and CYP97C. Thus, it was assumed that for the efficient flux of the pathway and the synthesis of lutein, the observed protein–protein interaction between the CYP97 enzymes may be critical, whereas HYD4 did not seem to interact with the CYP97s [99, 127]. It is worth noting that for the plastidial CYP97 subfamily modifications were not necessary to facilitate expression in *E. coli*. This can probably be explained by the similarity of the plant cell plastids with the prokaryotic progenitors. CYP97s seem to be weakly associated with the envelope membrane of the plastids and they are not predicted to have any transmembrane domain [99].

## *4.2 Yeast Systems*

Yeast has been used successfully for the expression of complex plant pathways including the P450s, in part due to the genomic tools available for this organism. Yeast has provided an alternative host for the functional characterisation of plant P450s recalcitrant to expression in bacterial systems [36, 129].

One of the most pronounced achievements of synthetic biology in high-value terpenoids in yeast is the synthesis of high titres of artemisinic acid with Paddon

and coworkers achieving titres of 25 g $L^{-1}$ [94]. To make this possible, multiple optimisation steps were required for efficient activity of the CYP71AV1 including the balancing of expression levels of CYP71AV1 in relation to the *A. annua* POR, and coexpression of an additional reductase, *A. annua* cytochrome *b5*. These changes increased the viability of the yeast cells, but the important developments that led to the high titres mentioned above were the discovery of the entire biosynthetic pathway of artemisinin including an *A. annua* alcohol dehydrogenase and artemisinic aldehyde dehydrogenase. The expression of these enzymes resulted in increased production of artemisinic acid from 2.4 to 8.1 g $L^{-1}$. Further improvements in fermentation tools and procedures then led to the final high titres [94]. These results show not only the remarkable potentials of synthetic biology and metabolic engineering for the synthesis of valuable compounds from microorganisms, but also provide a future strategy that can be applied to the production of other valuable terpenoids. A detailed description of this work can be found in the relevant chapter of this volume.

An interesting example demonstrating the potentials and also pitfalls of yeast bioengineering is the work regarding (+)-nootkatone synthesis. The development of several production platforms has been attempted using a range of hosts from bacteria to plants, yet to date the achieved titres of nootkatone are not sufficient for commercial applications [125]. Nootkatone can be produced from (+)-valencene, abundant in nature, through an enzymatic reaction assigned to a P450 enzyme. Due to the lack of a native plant enzyme at that time, several bacterial P450s were tested instead without success [37]. CYP71AV8 from chicory as well as the related tobacco CYP71D51v2 were expressed in yeast together with the Arabidopsis POR1 [17, 37]. In vitro results showed that both enzymes were able to catalyse the conversion of valencene to nootkatol efficiently but only to a minor degree to nootkatone. Although very small amounts of nootkatone were produced from yeast microsomes expressing CYP71D51v2, in vivo bioconversion (using the same yeast strain) of valencene and β-nootkatol to nootkatone proved to be much more efficient, yielding up to 3 mg $L^{-1}$ of nootkatone. However, this production was found subsequently to be independent of CYP71D51v2, and likely catalysed by a yeast endogenous enzyme. Although increasing the substrate concentration led to the decrease of the final product due to toxicity issues, 6 mg $L^{-1}$ of β-nootkatol and nootkatone could finally be produced from 200 mg $L^{-1}$ of valencene [37].

A P450 from the Solanaceae family resulted in more promising results. Henbane (*Hyoscyamus muticus*) CYP71D55 is an enzyme previously shown to oxidise various decalin-ring sesquiterpenoids, including valencene [117]. The methylotrophic *Pichia pastoris* was used as host and has been shown to be tolerant for heterologous expression of specific P450s [125]. A variant of CYP71D55, mutated in two amino acids, V482I and A484I, improved the catalytic efficiency of the enzyme for nootkatol biosynthesis up to fivefold, but nevertheless failed to convert it to nootkatone [117]. Despite this observed lack of function of CYP71D55, the *P. pastoris* strain expressing the valencene synthase and CYP71D55/POR accumulated low amounts of nootkatone. It was shown that an endogenous *P. pastoris* alcohol dehydrogenase was responsible for this conversion. Overexpression of this

enzyme, together with the remaining enzymes of the pathway (valencene synthase, CYP71D55, POR) and a truncated version of yeast HMG1 resulted in the production of 208 mg $L^{-1}$ of nootkatone in fed-batch cultures. This production titre reaches the range of commercially sustainable production in grams per litre, and awaits further strain improvements [125].

With identification of CYP706M1, the native P450 from cypress catalysing formation of nootkatone from valencene, a turning point was reached [18]. In contrast to previous assays, expression of CYP706M1 in yeast cells resulted in nearly quantitative oxidation of valencene to nootkatone. Microsomal assays of yeast cells expressing CYP706M1 resulted mainly in the synthesis of nootkatone (82 % of total products) whereas only minor amounts of *trans*-nootkatol were observed. This indicates that CYP706M1 efficiently performs a two-step oxidation. Despite the high specificity of CYP706M1, engineering of a yeast strain coexpressing the Arabidopsis POR1 and the cypress valencene synthase resulted in 144 μg $L^{-1}$ of nootkatone, whereas the control yeast strain lacking the P450 accumulated almost tenfold higher levels of valencene. This indicates that a bottleneck of the pathway is likely at the level of CYP706M1, making it a prime target for optimisation by enzyme engineering [18].

Triterpenoid saponins are a group of specialised metabolites found in many plant species and they are known for their diverse biological activities [4]. Because of their potential use as pharmaceutical compounds, large-scale biotechnological synthesis of these metabolites is desirable. Fukushima and coworkers showed that CYP716A12 and CYP93E2 are enzymes involved in the first step of the pathway to bioactive saponins, functionalising β-amyrin towards hemolytic and nonhemolytic sapogenins in the model legume plant *M. truncatula* [35]. Subsequently, the next step of oxidations catalysed by CYP72A68v2 and CYP72A61v2 was defined. In yeast strains coexpressing the *M. truncatula* P450s pairs CYP716A12/CYP72A68v2 and CYP93E2/CYP72A61v2 with the *Lotus japonicus* β-amyrin synthase and POR, gypsogenic acid and soyasapogenol B, respectively, were detected. Further, coexpression of novel combinations of P450s resulted in an array of novel terpenoids not yet detected in *M. truncatula*, including 4-epi-hederagenin, queretaroic acid, oleanolic acid, and β-amyrin derivatives oxidised at different carbon molecules [36]. Thus, it seems possible to combine P450 enzymes in yeast that do not co-occur in their native host for the synthesis of novel compounds with possible original biological functions.

One of the few examples of coexpression of multiple P450s in yeast aims at the synthesis of ginsenoside aglycons by Dai and coworkers [26]. Ginsenosides are the main bioactive phytochemicals of ginseng belonging to the triterpenoids class. The three basic aglycons of ginsenosides are protopanaxadiol, protopanaxatriol, and oleanolic acid. They are synthesised through cyclisation of 2,3-oxidosqualene by β-amyrin and dammarenediol-II synthase. The conversion of β-amyrin to oleanolic acid and of dammarenediol-II to protopanaxadiol and consequently to protopanaxatriol is catalysed by a number of previously identified P450 enzymes [35, 45]. The enzymes responsible for the synthesis of oleanolic acid, protopanaxadiol and protopanaxatriol were coexpressed in yeast by genomic integration together with enzymes from early steps of the pathway (a truncated HMG1,

squalene synthase, and squalene epoxidase from yeast). The origin of the oleanolic acid synthase (CYP716A12) was *M. truncatula* and the protopanaxadiol (CYP716A47) and protopanaxatriol (CYP716A53v2) synthases were derived from *P. ginseng*. After coupling with *Arabidopsis* POR1 the P450s were functional in yeast without further modification yielding 17.2 mg $L^{-1}$ protopanaxadiol, 15.9 mg $L^{-1}$ protopanaxatriol, and 21.4 mg $L^{-1}$ oleanolic acid in nonoptimised flask cultures [26]. These results may encourage further attempts at coexpression and assembly of pathways utilising P450s from multiple species.

### *4.3 Plant Systems: Nicotiana Benthamiana/Agrobacterium*

The use of plant hosts for the biotechnological production of terpenoids appears to be an obvious approach because the isoprene precursors, cofactors, and critical enzymes, such as P450-compatible PORs required for terpenoid biosynthesis, are endogenously present. Furthermore, the protein translation machinery, posttranslational modifications, cellular compartments, and the use of transit peptides for protein targeting, are conserved among plants, eliminating some of the problems encountered when attempting to express plant genes in microbial hosts. Yet, plant systems remain largely underexplored. Possible reasons for this are that plants present some specific complications as hosts, notably the long turnover times for generation of transgenic lines expressing new proteins, the risk for transgene-silencing, highly differentiated cell types, and endogenous terpenoid biochemistry. *N. benthamiana* is one of the few examples of a plant expression system successfully used at an industrial scale (although its major use is still the production of pharmaceutically active peptides). Prominent examples of molecules made using the transient *N. benthamiana/Agrobacterium* expression system are vaccines for the treatment of malaria, influenza, and Ebola [32, 63, 92]. An advantage of this system is the short time between infiltration and harvest of typically four to six days, as well as its industrial scalability [41], and distinct and controllable glycosylation machinery [19, 75]. Concerning terpenoid production, *N. benthamiana* is well established and the robustness of the system has facilitated both high-throughput testing of enzymes and accumulation of substantial amounts of terpenes [5, 13].

The sesquiterpene lactone costunolide is of industrial interest due to inherent biological activities and as starting material for a broad range of high-value compounds such as the pharmaceutical parthenolide found in the medicinal plant feverfew (Tanacetum parthenium; [71]). The pathway to the precursor molecule, costunolide, is known and involves the formation of germacrene A by germacrene A synthase, oxidation at C-12 to germacrene A acid by CYP71AV (germacrene A oxidase; [70]), and then hydroxylation at C-6 by CYP71BL subfamily members. The product of this final reaction autocatalytically rearranges to costunolide [27, 47, 70]. With this pathway elucidated, transient expression in *N. benthamiana* as a potential production host was approached [70]. To optimise production of the first intermediate, several germacrene A synthases were evaluated and the most active

enzyme from feverfew was subsequently targeted to the mitochondria using the targeting sequence of cytochrome oxidase subunit IV (CoxIV) leading to substantially improved production of germacrene A. Coexpression of both P450s, and the mitochondrially targeted germacrene A synthase, all driven by the strong constitutive Rubisco-promoter, resulted in production of up to 60 ng $g^{-1}$ FW of costunolide. Nontargeted metabolomics showed the accumulation of considerable amounts of glutathione and cysteine conjugates of costunolide indicating that potential pathways for detoxification are active in planta. Although conjugation could present a potential problem in synthetic biology, it was on the other hand suggested as an opportunity to improve production and storage of terpenoids due to increased water solubility [70]. In plants, conjugation has the dual effect of masking reactive functional groups and increasing solubility, possibly targeting the molecules for further metabolisation and/or sequestration in the vacuole.

With formation of costunolide established, Liu and coworkers approached the final step, oxidation of costunolide to parthenolide. Coexpression of the established pathway together with feverfew CYP71CA1 and a soluble *Arabidopsis* HMG-CoA reductase, to increase substrate availability, yielded 2.05 ng $g^{-1}$ FW of free parthenolide. Again, both parthenolide–cysteine and parthenolide–glutathione were found to constitute more than 90 % of the total amount of sesquiterpenoid produced [71]. A similar observation of effective conjugation of produced terpenoids was made when the model pathway to artemisinic acid was introduced into *N. benthamiana.* Here, predominant accumulation of approximately 40 ng $g^{-1}$ FW artemisinic acid-12-β-diglycoside accompanied by absence of the nonconjugated terpenoid indicated efficient conversion by the endogenous conjugation machinery [123]. When reconstructing parts of the artemisinin pathway, LC-MS analysis of extracts revealed the presence of at least 19 conjugated forms of these functionalised sesquiterpenoids. Compounds displayed varying levels of glycosylation as well as conjugation with malonate and glutathione [120]. For construction of multistep pathways, it was suggested that the problem of conjugation could be alleviated through efficient and balanced expression of pathway enzymes to facilitate fast channelling of activated metabolic intermediates to downstream products [100].

Clearly, the conjugation of mono- and sesquiterpenoids oxidatively functionalised with hydroxyl or carboxyl groups by plant hosts needs to be considered in biotechnological production. However, in the few recent examples of *N. benthamiana*-based functional testing of P450 containing pathways to di- and triterpenoids, such as functional expression of the *Lotus japonicus* CYP71D353 which catalyses the formation of 20-hydroxybetulinic acid in a sequential three-step oxidation of 20-hydroxylupeol [60], no evidence was found indicating that this phenomenon may not play a significant role for larger and nonvolatile terpenoids.

Even though *N. benthamiana* is now routinely used for functional analysis of terpenoid biosynthetic pathways, including P450s, these few examples above indicate a promising future potential for targeted production of terpenoids. One limitation, though, which will need to be addressed in the future, is the production capacity of plants, with solutions possibly in storage of produced terpenoids in

specialised structures such as glandular trichomes and regulation of the pathways supplying the universal C5 building blocks.

### *4.4 Plant Systems: Physcomitrella patens*

The moss *Physcomitrella patens* represents a plant system uniquely suited to high-value terpenoid production due to a natural tolerance to accumulation of terpenoids, a simple endogenous terpenoid profile, and established genetic tools for genome editing by homologous recombination [6, 133]. In contrast to higher plant species, *P. patens* lacks gibberellin diterpene phytohormones but accumulates *ent*-kaurene intermediates of the gibberellin pathway. To date, there are at least two examples reported of the successful biosynthesis of nonnative terpenes by this moss. Constitutive expression of the taxadiene diterpene synthase from Pacific yew (*Taxus brevifolia*) resulted in the accumulation of up to 0.05 % tissue fresh weight of taxa-4(5),11(12)-diene, without significantly changing the accumulation or pattern of endogenous diterpenoids [3]. The *P. patens* genome encodes only a single diterpene synthase (copalyl-diphosphate/*ent*-kaurene synthase), which can be disrupted without phenotypic consequences in the relevant life stages. This was exploited for stable introduction of the single-step terpenoid biosynthetic pathways to the sesquiterpenoids patchoulol and santalene [133]. Although there are currently no published examples of the heterologous expression of functionally characterised P450s in *P. patens*, it is expected that this plant host will provide the same benefits as *N. benthamiana* such as the presence of the required subcellular compartments, partner enzymes, and cofactors, but with the added benefits of homologous recombination, short regeneration times, and scalable production in bioreactors.

## 5 Outlook and Perspectives

Significant technological advances have increased our abilities to discover, characterise, and manipulate P450s and this has led to breakthroughs in the biotechnological production of complex terpenoids (e.g., the production of artemisinic acid in yeast [94]). It is clear that major challenges remain.

The ability to combine enzymes from different terpenoid pathways and, to a degree, design new P450s [22] opens up the possibility of generating new to nature compounds with novel or specific bioactivities. Several groups have already demonstrated that this approach is feasible with novel triterpenes being generated from nonnatural combinations of plant-derived triterpene synthases and P450s expressed in yeast [26, 36, 78].

As photosynthesis in the chloroplast provides energy in the form of ATP, reducing equivalents, and a carbon source for isoprenoid $C_5$ building blocks, photosynthetic production systems may offer unique opportunities for development

of new efficient P450-based pathways. Recent work has shown that it is possible to relocate P450s to the thylakoid membrane of the chloroplast and divert electrons directly from active photosynthesis to drive P450-mediated reactions, without the requirement of POR enzymes [64, 89]. Light-driven activity of P450s may therefore relieve current limitations of biotechnological applications, in particular the requirement of the membrane-anchored POR and stoichiometric amounts of NADPH. Consequently, this technology has a high potential for engineering of pathways already native to the plastids (i.e., mono- and diterpenoids), or those engineered for plastid localisation (i.e., sesqui- and triterpenoids; e.g., [105, 126]). Furthermore, recent progress in the development of tools for plastid genome transformation enables high-level expression of heterologous enzymes in chloroplasts due to the lack of transgene silencing and epigenetic mechanisms [10]. However, despite the encouraging successes and the inherent potential, there are currently no reports of engineering plastids or their autonomous equivalents, cyanobacteria such as *Synechocystis* or *Synechococcus* which harbour the pathways for $C_5$ isoprene building blocks [31], with P450s involved in terpenoid biosynthesis.

As it has become increasingly feasible to identify pathways of specialised metabolism, research focus is shifting towards understanding how the enzymes of these pathways interact and how these interactions affect pathway activity. In the case of terpenoid metabolism, this question relates to how the various involved P450s, PORs, and other ER or cytosolic enzymes, such as glycosyl-transferases and acyl-transferases, interact. There is growing evidence that the enzymes involved in a given pathway form dynamic multienzyme complexes (termed 'metabolons'; [54]). Metabolon formation may enable channelling of substrates, thought to result in more efficient and directed pathways, as well as provide an important means to regulate pathway activity [65]. In the case of plant-specialised metabolism, metabolon formation could be an important mechanism to reduce exposure of potentially toxic or labile intermediates to the cellular environment [65]. Currently our understanding of the factors enabling and controlling the formation and dissociation of metabolons is quite limited, but knowledge in this field is expected to expand with available new technologies. If metabolon formation does play an important role in pathway channelling, understanding how to control this phenomenon could provide significant efficiency gains for biotechnological production of terpenoids.

**Acknowledgments** Financial support was provided by the UNIK [Universitetsforskningens Investeringskapital (Investment Capital for University Research)] Center for Synthetic Biology "bioSYNergy" at the University of Copenhagen, funded by the Research Initiative of the Danish Ministry of Science, Technology and Innovation (BH), and the Novo Nordisk Foundation (BH). Additional support was provided by a personal postdoctoral stipend awarded by the FP7 PEOPLE MARIE CURIE ACTIONS Intra-European Fellowships (IP).

## A.1 6 Appendix

See Table 2.

**Table 2** P450s involved in terpenoid biosynthesis identified during 2013–2014

| Clan | Subfamily | Metabolism | Class of metabolites | Function | Plant species | References |
|---|---|---|---|---|---|---|
| CYP51 | CYP51H | Specialised | Triterpenoid | Triterpene avenacin biosynthesis (epoxydase); As CYP51H10 is a multifunctional P450 capable of modifying both the C and D rings of the pentacyclic triterpene scaffold to give 12,13β-epoxy-3β,16β-dihydroxy-oleanane (12,13β-epoxy-16β-hydroxy-β-amyrin) | Monocotelydons; oat (*Avena* spp.) | Geisler et al. [38] |
| CYP71 | CYP71CA1 | Specialised | Sesquiterpenoid | Epoxidation of costunolide to parthenolide | Feverfew (*Tanacetum parthenium*) | Liu et al. [71] |
| | CYP71D | Specialised | Mixed, including: sesquiterpenoid, diterpenoid, indole alkaloid, flavonoid | 5-epi-aristolochene and 1-deoxycapsidiol to capsidiol in vitro, cembratrienol hydroxylase, limonene 3- and 6-hydroxylase, menthol hydroxylase, thymol and carvacrol, two-step production of (+)-trans-piperitenol, successive reactions at carbon 2 (C-2) of premnaspirodiene primarily to nootkatol in vitro, cembratriene-ol do CBT-diol, flavonoid hydroxylase, tabersonine hydroxylase | Tobacco (*Nicotiana tabacum*), mint, *Catharantus roseus*, *Hyoscyamus muticus*, *Ocimum* spp. | Gavira et al. [37] |
| | CYP76AH | Specialised | Diterpenoid | CYP76AH hydroxylate miltiradiene/abietatriene | Lamiaceae | Guo et al. [40], Zi et al. [134] |
| | CYP76F | Specialised | Sesquiterpenoid | A series of nine P450 was functionally characterized using in vitro and yeast in vivo assays and shown to encode santalene/bergamotene oxidases and bergamotene oxidases | Sandalwoold | Diaz-Chavez et al. [29] |
| | CYP93E | Specialised | Triterpenoid | C-24 hydroxylation of β-amyrin: hydroxylates triterpenes (β-amyrin or sophoradiol) to olean-12-ene-3β,24-diol and soyasapogenol B in both soybean and licorice (G. uralensis); | soybean (*Glycine max* L.); licorice (Glycyrrhiza); Medicago truncatula; | Fukushima et al. [36] |
| | CYP706M | Specialised | Sesquiterpenoid | Valencene oxidase producing trans-nootkatol and (+)-nootkatone in yeast | Alaska cedar (*Callitropsis nootkatensis* ) | Cankar et al. [18] |
| | CYP726A | Specialised | Diterpenoid | Oxidation of macrocyclic diterpenes casbene, cembrene | Euphorbiaceae | King et al. [58] |

(continued)

**Table 2** (continued)

| Clan | Subfamily | Metabolism | Class of metabolites | Function | Plant species | References |
|---|---|---|---|---|---|---|
| CYP72 | CYP714A | General | Diterpenoid phytohormone GA | Inactivation of early GA intermediates; CYP714A1 catalyzes conversion of GA12 to 16-carboxylated GA12 (16-carboxy-16β,17-dihydro GA12), with GA12 12α-hydroxy GA12 (GA111) was produced; CYP714A2 catalyzed conversion of *ent*-kaurenoic acid into steviol (*ent*-13-hydroxy kaurenoic acid) | Arabidopsis | Nomura et al. [91] |
| | CYP714B | General | Diterpenoid phytohormone GA | Gibberellin 13-oxidases that reduce gibberellin activity; Transgenic Arabidopsis plants that overexpress CYP714B1 or CYP714B2 show semidwarfism | Rice (*Oryza sativa* ) | Magome et al. [74] |
| | CYP72A | Specialised | Monoterpene indole alkaloid, triterpene saponins | Secologanin synthesis, catalyzing the conversion of loganin to secologanin and from 7-dexyloganic acid to loganic acid; ring-opening of cyclopentane ring through C–C cleavage; CYP72A154 catalyzed three sequential oxidation steps at C-30 of 11-oxo-β-amyrin to glycyrrhetinic acid; CYP72A63 catalyzes C-30 oxidation of β-amyrin | *Catharanthus roseus*, licorice (*Glycyrrhiza*), *Medicago truncatula* | Fukushima et al. [36], Salim et al. [104] |
| CYP85 | CYP716A | Specialised | Triterpenoid | Multifunctional with β-amyrin 28-oxidase, α-amyrin 28-oxidase, and lupeol 28-oxidase activities; oxidation of β-amyrin and erythrodiol at C-28 position, yielding oleanolic acid. CYP716A47 and CYP716A53v2 catalyze the formation of dammarene-type triterpenesd in ginseng (CYP716A47 catalyzes formation of protopanaxadiol from dammarenediol-II and CYP716A53v2 catalyzes formation of protopanaxatriol from protopanaxadiol) | *Medicago truncatula*, grape, ginseng (Panax ginseng), genus Bupleurum | Fukushima et al. [36], Moses et al. [78] |
| CYP97 | CYP97C | Specialised | Acyclic diterpenoid | Biosynthesis of plaunotol (18-OH geranylgeraniol) | *Croton stellatopilosus* | Sintupachee et al. [115] |

See Hamberger and Bak [42] for a comprehensive list

## References


1. Ajikumar PK, Xiao W-H, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G (2010) Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. Science 330(6000):70–74
2. Alonso-Gutierrez J, Chan R, Batth TS, Adams PD, Keasling JD, Petzold CJ, Lee TS (2013) Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. Metab Eng 19:33–41
3. Anterola A, Shanle E, Perroud PF, Quatrano R (2009) Production of taxa-4(5),11(12)-diene by transgenic *Physcomitrella patens*. Transgenic Res 18(4):655–660
4. Augustin JM, Kuzina V, Andersen SB, Bak S (2011) Molecular activities, biosynthesis and evolution of triterpenoid saponins. Phytochemistry 72(6):435–457
5. Bach SS, Bassard JE, Andersen-Ranberg J, Moldrup ME, Simonsen HT, Hamberger B (2014) High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in *Nicotiana benthamiana*. Methods Mol Biol 1153:245–255
6. Bach SS, King BC, Zhan X, Simonsen HT, Hamberger B (2014) Heterologous stable expression of terpenoid biosynthetic genes using the moss *Physcomitrella patens*. Methods Mol Biol 1153:257–271
7. Bak S, Kahn RA, Nielsen HL, Moller BL, Halkier BA (1998) Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. Plant Mol Biol 36(3):393–405
8. Barnes HJ, Arlotto MP, Waterman MR (1991) Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. Proc Natl Acad Sci 88(13):5597–5601
9. Bertea CM, Freije JR, van der Woude H, Verstappen FW, Perk L, Marquez V, De Kraker JW, Posthumus MA, Jansen BJ, de Groot A, Franssen MC, Bouwmeester HJ (2005) Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. Planta Med 71(1):40–47
10. Bock R (2014) Genetic engineering of the chloroplast: novel tools and new applications. Curr Opin Biotechnol 26:7–13
11. Bohlmann J, Keeling CI (2008) Terpenoid biomaterials. Plant J Cell Mol Biol 54(4):656–669
12. Boutanaev AM, Moses T, Zi J, Nelson DR, Mugford ST, Peters RJ, Osbourn A (2014) Investigation of terpene diversification across multiple sequenced plant genomes. Proc Natl Acad Sci
13. Bruckner K, Tissier A (2013) High-level diterpene production by transient expression in *Nicotiana benthamiana*. Plant Methods 9(1):46
14. Butler MS (2008) Natural products to drugs: natural product-derived compounds in clinical trials. Nat Prod Rep 25(3):475–516
15. Butler CF, Peet C, McLean KJ, Baynham MT, Blankley RT, Fisher K, Rigby SE, Leys D, Voice MW, Munro AW (2014) Human P450-like oxidation of diverse proton pump inhibitor drugs by 'gatekeeper' mutants of flavocytochrome P450 BM3. Biochem J 460(2):247–259. LID—210.1042/BJ20140030 [doi]
16. Caniard A, Zerbe P, Legrand S, Cohade A, Valot N, Magnard JL, Bohlmann J, Legendre L (2012) Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. BMC Plant Biol 12:119
17. Cankar K, van Houwelingen A, Bosch D, Sonke T, Bouwmeester H, Beekwilder J (2011) A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. FEBS Lett 585(1):178–182

18. Cankar K, van Houwelingen A, Goedbloed M, Renirie R, de Jong RM, Bouwmeester H, Bosch D, Sonke T, Beekwilder J (2014) Valencene oxidase CYP706M1 from Alaska cedar (*Callitropsis nootkatensis*). FEBS Lett 588(6):1001–1007
19. Castilho A, Bohorova N, Grass J, Bohorov O, Zeitlin L, Whaley K, Altmann F, Steinkellner H (2011) Rapid high yield production of different glycoforms of Ebola virus monoclonal antibody. PLoS ONE 6(10):e26040
20. Chang MCY, Eachus RA, Trieu W, Ro D-K, Keasling JD (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. Nat Chem Biol 3(5):274–277
21. Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66(1):212–229
22. Coelho PS, Brustad EM, Kannan A, Arnold FH (2013) Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. Science 339(6117):307–310
23. Cragg GM (1998) Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development. Med Res Rev 18(5):315–331
24. Croteau R, Ketchum RB, Long R, Kaspera R, Wildung M (2006) Taxol biosynthesis and molecular genetics. Phytochem Rev 5(1):75–97
25. Cyr A, Wilderman PR, Determan M, Peters RJ (2007) A modular approach for facile biosynthesis of labdane-related diterpenes. J Am Chem Soc 129(21):6684–6685
26. Dai Z, Wang B, Liu Y, Shi M, Wang D, Zhang X, Liu T, Huang L, Zhang X (2014) Producing aglycons of ginsenosides in bakers' yeast. Sci Rep 4:3698. doi:10.1038/srep03698
27. de Kraker JW, Franssen MC, Joerink M, de Groot A, Bouwmeester HJ (2002) Biosynthesis of costunolide, dihydrocostunolide, and leucodin. Demonstration of cytochrome p450-catalyzed formation of the lactone ring present in sesquiterpene lactones of chicory. Plant Physiol 129(1):257–268
28. DeJong JM, Liu Y, Bollon AP, Long RM, Jennewein S, Williams D, Croteau RB (2006) Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. Biotechnol Bioeng 93(2):212–224
29. Diaz-Chavez ML, Moniodis J, Madilao LL, Jancsik S, Keeling CI, Barbour EL, Ghisalberti EL, Plummer JA, Jones CG, Bohlmann J (2013) Biosynthesis of sandalwood oil: *Santalum album* CYP76F cytochromes P450 produce santalols and bergamotol. PLoS ONE 8(9): e75053
30. Dietrich JA, Yoshikuni Y, Fisher KJ, Woolard FX, Ockey D, McPhee DJ, Renninger NS, Chang MCY, Baker D, Keasling JD (2009) A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450BM3. ACS Chem Biol 4(4):261–267
31. Englund E, Pattanaik B, Ubhayasekera SJ, Stensjo K, Bergquist J, Lindberg P (2014) Production of squalene in *Synechocystis* sp. PCC 6803. PLoS ONE 9(3):e90270
32. Feller T, Thom P, Koch N, Spiegel H, Addai-Mensah O, Fischer R, Reimann A, Pradel G, Fendel R, Schillberg S, Scheuermayer M, Schinkel H (2013) Plant-based production of recombinant *Plasmodium* surface protein pf38 and evaluation of its potential as a vaccine candidate. PLoS ONE 8(11):e79920
33. Feyereisen R (2011) Arthropod CYPomes illustrate the tempo and mode in P450 evolution. Biochim Biophys Acta 1814(1):19–28
34. Frense D (2007) Taxanes: perspectives for biotechnological production. Appl Microbiol Biotechnol 73(6):1233–1240
35. Fukushima EO, Seki H, Ohyama K, Ono E, Umemoto N, Mizutani M, Saito K, Muranaka T (2011) CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. Plant Cell Physiol 52(12):2050–2061
36. Fukushima EO, Seki H, Sawai S, Suzuki M, Ohyama K, Saito K, Muranaka T (2013) Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. Plant Cell Physiol 54(5):740–749

37. Gavira C, Höfer R, Lesot A, Lambert F, Zucca J, Werck-Reichhart D (2013) Challenges and pitfalls of P450-dependent (+)-valencene bioconversion by *Saccharomyces cerevisiae*. Metab Eng 18:25–35
38. Geisler K, Hughes RK, Sainsbury F, Lomonossoff GP, Rejzek M, Fairhurst S, Olsen C-E, Motawia MS, Melton RE, Hemmings AM, Bak S, Osbourn A (2013) Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. Proc Natl Acad Sci 110(35):E3360–E3367
39. Gongora-Castillo E, Childs KL, Fedewa G, Hamilton JP, Liscombe DK, Magallanes-Lundback M, Mandadi KK, Nims E, Runguphan W, Vaillancourt B, Varbanova-Herde M, Dellapenna D, McKnight TD, O'Connor S, Buell CR (2012) Development of transcriptomic resources for interrogating the biosynthesis of monoterpene indole alkaloids in medicinal plant species. PLoS ONE 7(12):e52506
40. Guo J et al (2013) PNAS 110(29):12108–12113
41. Hahn S, Giritch A, Bartels D, Bortesi L, Gleba Y (2014) A novel and fully scalable *Agrobacterium* spray-based process for manufacturing cellulases and other cost-sensitive proteins in plants. Plant Biotechnol J. doi:10.1111/pbi.12299 Early online version [Epub ahead of print]
42. Hamberger B, Bak S (2013) Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Philos Trans R Soc B: Biol Sci 368(1612):20120426. doi:10.1098/rstb.2012.0426
43. Hamberger B, Ohnishi T, Hamberger B, Séguin A, Bohlmann J (2011) Evolution of diterpene metabolism: sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. Plant Physiol 157(4):1677–1695
44. Han JY, Kim HJ, Kwon YS, Choi YE (2011) The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. Plant Cell Physiol 52(12):2062–2073
45. Han J-Y, Hwang H-S, Choi S-W, Kim H-J, Choi Y-E (2012) Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax Ginseng*. Plant Cell Physiol 53(9):1535–1545
46. Hardy K, Buckley S, Collins MJ, Estalrrich A, Brothwell D, Copeland L, Garcia-Tabernero A, Garcia-Vargas S, de la Rasilla M, Lalueza-Fox C, Huguet R, Bastir M, Santamaria D, Madella M, Wilson J, Cortes AF, Rosas A (2012) Neanderthal medics? Evidence for food, cooking, and medicinal plants entrapped in dental calculus. Naturwissenschaften 99(8):617–626
47. Ikezawa N, Göpfert JC, Nguyen DT, Kim S-U, O'Maille PE, Spring O, Ro D-K (2011) Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. J Biol Chem 286(24):21601–21611
48. Jackson CJ, Lamb DC, Marczylo TH, Warrilow AGS, Manning NJ, Lowe DJ, Kelly DE, Kelly SL (2002) A novel sterol 14α-demethylase/ferredoxin fusion protein (MCCYP51FX) from *Methylococcus capsulatus* represents a new class of the cytochrome P450 superfamily. J Biol Chem 277(49):46959–46965
49. Jennewein S, Rithner CD, Williams RM, Croteau RB (2001) Taxol biosynthesis: taxane 13α-hydroxylase is a cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci 98 (24):13595–13600
50. Jennewein S, Rithner CD, Williams RM, Croteau R (2003) Taxoid metabolism: taxoid 14β-hydroxylase is a cytochrome P450-dependent monooxygenase. Arch Biochem Biophys 413(2):262–270
51. Jennewein S, Long RM, Williams RM, Croteau R (2004) Cytochrome P450 taxadiene 5α-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. Chem Biol 11(3):379–387
52. Jensen K, Møller BL (2010) Plant NADPH-cytochrome P450 oxidoreductases. Phytochemistry 71(2–3):132–141
53. Jiang M, Stephanopoulos G, Pfeifer BA (2012) Toward biosynthetic design and implementation of *Escherichia coli*-derived paclitaxel and other heterologous polyisoprene compounds. Appl Environ Microbiol 78(8):2497–2504

54. Jorgensen K, Rasmussen AV, Morant M, Nielsen AH, Bjarnholt N, Zagrobelny M, Bak S, Moller BL (2005) Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Curr Opin Plant Biol 8(3):280–291
55. Kaspera R, Croteau R (2006) Cytochrome P450 oxygenases of taxol biosynthesis. Phytochem Rev 5(2–3):433–444
56. Kato-Noguchi H, Peters RJ (2013) The role of momilactones in rice allelopathy. J Chem Ecol 39(2):175–185
57. Kim J-E, Cheng KM, Craft NE, Hamberger B, Douglas CJ (2010) Over-expression of *Arabidopsis thaliana* carotenoid hydroxylases individually and in combination with a β-carotene ketolase provides insight into in vivo functions. Phytochemistry 71(2–3):168–178
58. King AJ, Brown GD, Gilday AD, Larson TR, Graham IA (2014) Production of bioactive diterpenoids in the euphorbiaceae depends on evolutionarily conserved gene clusters. Plant Cell 26(8):3286–3298
59. Kirby J, Keasling JD (2009) Biosynthesis of plant isoprenoids: perspectives for microbial engineering. Annu Rev Plant Biol 60(1):335–355
60. Krokida A, Delis C, Geisler K, Garagounis C, Tsikou D, Pena-Rodriguez LM, Katsarou D, Field B, Osbourn AE, Papadopoulou KK (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol 200 (3):675–690
61. Kunii M, Kitahama Y, Fukushima EO, Seki H, Muranaka T, Yoshida Y, Aoyama Y (2012) b-Amyrin oxidation by oat CYP51H10 expressed heterologously in yeast cells: the first example of CYP51-dependent metabolism other than the 14-demethylation of sterol precursors. Biol Pharm Bull 35(5):801–804
62. Lamb DC, Waterman MR (2013) Unusual properties of the cytochrome P450 superfamily. Philos Trans R Soc B: Biol Sci 368(1612):20120434. doi:10.1098/rstb.2012.0434
63. Landry N, Pillet S, Favre D, Poulin JF, Trepanier S, Yassine-Diab B, Ward BJ (2014) Influenza virus-like particle vaccines made in *Nicotiana benthamiana* elicit durable, poly-functional and cross-reactive T cell responses to influenza HA antigens. Clin Immunol 154 (2):164–177
64. Lassen LM, Nielsen AZ, Olsen CE, Bialek W, Jensen K, Moller BL, Jensen PE (2014) Anchoring a plant cytochrome P450 via PsaM to the thylakoids in *Synechococcus* sp. PCC 7002: evidence for light-driven biosynthesis. PLoS ONE 9(7):e102184
65. Laursen T, Møller BL, Bassard J-E (2014) Plasticity of specialized metabolism as mediated by dynamic metabolons. Trends Plant Sci. doi:10.1016/j.tplants.2014.11.002 [Epub ahead of print]
66. Leonard E, Koffas MAG (2007) Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. Appl Environ Microbiol 73(22):7246–7251
67. Leroi-Gourhan A (1975) The flowers found with Shanidar IV, a neanderthal burial in Iraq. Science 190(4214):562–564
68. Li Y, Pfeifer BA (2014) Heterologous production of plant-derived isoprenoid products in microbes and the application of metabolic engineering and synthetic biology. Curr Opin Plant Biol 19:8–13
69. Liu T, Khosla C (2010) A balancing act for taxol precursor pathways in *E. coli*. Science 330 (6000):44–45
70. Liu Q, Majdi M, Cankar K, Goedbloed M, Charnikhova T, Verstappen FW, de Vos RC, Beekwilder J, van der Krol S, Bouwmeester HJ (2011) Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. PLoS ONE 6(8):e23255
71. Liu Q, Manzano D, Tanic N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S, Bouwmeester H (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. Metab Eng 23:145–153
72. Lovering F (2013) Escape from Flatland 2: complexity and promiscuity. Med Chem Comm 4 (3):515–519
73. Lovering F, Bikker J, Humblet C (2009) Escape from Flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52(21):6752–6756
74. Magome H et al (2013) PNAS 110(5):1947–1952

75. Mamedov T, Ghosh A, Jones RM, Mett V, Farrance CE, Musiychuk K, Horsey A, Yusibov V (2012) Production of non-glycosylated recombinant proteins in *Nicotiana benthamiana* plants by co-expressing bacterial PNGase F. Plant Biotechnol J 10(7):773–782
76. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437(7057):376–380
77. Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Pollier J, Woittiez L, van der Krol S, Lugan R, Ilc T, Verpoorte R, Oksman-Caldentey KM, Martinoia E, Bouwmeester H, Goossens A, Memelink J, Werck-Reichhart D (2014) The seco-iridoid pathway from *Catharanthus roseus*. Nat Commun 5:3606
78. Moses T, Pollier J, Almagro L, Buyst D, Van Montagu M, Pedreno MA, Martins JC, Thevelein JM, Goossens A (2014) Combinatorial biosynthesis of sapogenins and saponins in Saccharomyces cerevisiae using a C-16alpha hydroxylase from Bupleurum falcatum. Proc Natl Acad Sci U S A 111(4):1634–1639
79. Moses T, Pollier J, Almagro L, Buyst D, Van Montagu M, Pedreño MA, Martins JC, Thevelein JM, Goossens A (2014) Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16α hydroxylase from *Bupleurum falcatum*. Proc Natl Acad Sci 111(4):1634–1639
80. Moses T, Pollier J, Faizal A, Apers S, Pieters L, Thevelein JM, Geelen D, Goossens A (2014) Unravelling the triterpenoid saponin biosynthesis of the african shrub *Maesa lanceolata*. Mol Plant 8(1):122−135. doi:10.1016/j.molp.2014.11.004
81. Munro AW, Leys DG, McLean KJ, Marshall KR, Ost TWB, Daff S, Miles CS, Chapman SK, Lysek DA, Moser CC, Page CC, Dutton PL (2002) P450 BM3: the very model of a modern flavocytochrome. Trends Biochem Sci 27(5):250–257
82. Naoumkina MA, Modolo LV, Huhman DV, Urbanczyk-Wochniak E, Tang Y, Sumner LW, Dixon RA (2010) Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell 22(3):850–866
83. Nelson DR (1998) Metazoan cytochrome P450 evolution. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 121(1–3):15–22
84. Nelson DR (1999) Cytochrome P450 and the individuality of species. Arch Biochem Biophys 369(1):1–10
85. Nelson D, Werck-Reichhart D (2011) A P450-centric view of plant evolution. Plant J 66 (1):194–211
86. Nelson DR, Koymans L, Kamataki T, Stegeman J, Feyereisen R, Waxman D, Waterman M, Gotoh O, Coon M, Estabrook R, Gunsalus I, Nebert D (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 6 (1):1–42
87. Nelson DR, Schuler MA, Paquette SM, Werck-Reichhart D, Bak S (2004) Comparative genomics of rice and arabidopsis. analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. Plant Physiol 135(2):756–772
88. Newman DJ, Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod 75(3):311–335
89. Nielsen AZ, Ziersen B, Jensen K, Lassen LM, Olsen CE, Møller BL, Jensen PE (2013) Redirecting photosynthetic reducing power toward bioactive natural product synthesis. ACS Synth Biol 2(6):308–315
90. Noble MA, Miles CS, Chapman SK, Lysek DA, MacKay AC, Reid GA, Hanzlik RP, Munro AW (1999) Roles of key active-site residues in flavocytochrome P450 BM3. Biochem J 339:371–379

91. Nomura T, Magome H, Hanada A, Takeda-Kamiya N, Mander LN, Kamiya Y, Yamaguchi S (2013) Functional analysis of arabidopsis CYP714A1 and CYP714A2 reveals that they are distinct gibberellin modification enzymes. Plant Cell Physiol 54(11):1837–1851
92. Olinger GG Jr, Pettitt J, Kim D, Working C, Bohorov O, Bratcher B, Hiatt E, Hume SD, Johnson AK, Morton J, Pauly M, Whaley KJ, Lear CM, Biggins JE, Scully C, Hensley L, Zeitlin L (2012) Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques. Proc Natl Acad Sci U S A 109 (44):18030–18035
93. Paddon CJ, Keasling JD (2014) Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nat Rev Micro 12(5):355–367
94. Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Jiang H, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496 (7446):528–532
95. Panella NA, Dolan MC, Karchesy JJ, Xiong Y, Peralta-Cruz J, Khasawneh M, Montenieri JA, Maupin GO (2005) Use of novel compounds for pest control: insecticidal and acaricidal activity of essential oil components from heartwood of Alaska yellow cedar. J Med Entomol 42(3):352–358
96. Papadopoulou K, Melton RE, Leggett M, Daniels MJ, Osbourn AE (1999) Compromised disease resistance in saponin-deficient plants. Proc Natl Acad Sci U S A 96(22): 12923–12928
97. Peters RJ (2006) Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. Phytochemistry 67(21): 2307–2317
98. Qi X, Bakht S, Leggett M, Maxwell C, Melton R, Osbourn A (2004) A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. Proc Natl Acad Sci U S A 101(21):8233–8238
99. Quinlan RF, Shumskaya M, Bradbury LMT, Beltrán J, Ma C, Kennelly EJ, Wurtzel ET (2012) Synergistic interactions between carotene ring hydroxylases drive lutein formation in plant carotenoid biosynthesis. Plant Physiol 160(1):204–214
100. Renault H, Bassard JE, Hamberger B, Werck-Reichhart D (2014) Cytochrome P450-mediated metabolic engineering: current progress and future challenges. Curr Opin Plant Biol 19:27–34
101. Ro D-K, Arimura G-I, Lau SYW, Piers E, Bohlmann J (2005) Loblolly pine abietadienol/ abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. Proc Natl Acad Sci 102(22):8060–8065
102. Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440(7086):940–943
103. Saklani A, Kutty SK (2008) Plant-derived compounds in clinical trials. Drug Discov Today 13(3–4):161–171
104. Salim V, Yu F, Altarejos J, De Luca V (2013) Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis. Plant J 76(5):754–765
105. Saxena B, Subramaniyan M, Malhotra K, Bhavesh NS, Potlakayala SD, Kumar S (2014) Metabolic engineering of chloroplasts for artemisinic acid biosynthesis and impact on plant growth. J Biosci 39(1):33–41

106. Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L (2012) Toward a biosynthetic route to sclareol and amber odorants. J Am Chem Soc 134(46):18900–18903
107. Schmelz EA, Huffaker A, Sims JW, Christensen SA, Lu X, Okada K, Peters RJ (2014) Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. Plant J 79(4):659–678
108. Schoendorf A, Rithner CD, Williams RM, Croteau RB (2001) Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from taxus and functional expression in yeast. Proc Natl Acad Sci 98(4):1501–1506
109. Schuler MA, Werck-Reichhart D (2003) Functional genomics of P450s. Annu Rev Plant Biol 54(1):629–667
110. Schuler M, Duan H, Bilgin M, Ali S (2006) Arabidopsis cytochrome P450s through the looking glass: a window on plant biochemistry. Phytochem Rev 5(2–3):205–237
111. Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T (2008) Licorice beta-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. Proc Natl Acad Sci U S A 105 (37):14204–14209
112. Seki H, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, Fukushima EO, Akashi T, Aoki T, Saito K, Muranaka T (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. Plant Cell 23 (11):4112–4123
113. Sezutsu H, Le Goff G, Feyereisen R (2013) Origins of P450 diversity. Philos Trans R Soc B: Biol Sci 368(1612):20120428. doi:10.1098/rstb.2012.0428
114. Sharon-Asa L, Shalit M, Frydman A, Bar E, Holland D, Or E, Lavi U, Lewinsohn E, Eyal Y (2003) Citrus fruit flavor and aroma biosynthesis: isolation, functional characterization, and developmental regulation of Cstps1, a key gene in the production of the sesquiterpene aroma compound valencene. Plant J 36(5):664–674
115. Sintupachee S et al (2014) Plant Sci 229:131–141
116. Sullivan R, Behncke I, Purushotham A (2010) Why do we love medicines so much? An evolutionary perspective on the human love of pills, potions and placebo. EMBO Rep 11 (8):572–578
117. Takahashi S, Yeo Y-S, Zhao Y, O'Maille PE, Greenhagen BT, Noel JP, Coates RM, Chappell J (2007) Functional characterization of premnaspirodiene oxygenase, a cytochrome P450 catalyzing regio- and stereo-specific hydroxylations of diverse sesquiterpene substrates. J Biol Chem 282(43):31744–31754
118. Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS (2006) *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. FEBS Lett 580 (5):1411–1416
119. Tian L, Musetti V, Kim J, Magallanes-Lundback M, DellaPenna D (2004) The arabidopsis LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid ε-ring hydroxylation activity. Proc Natl Acad Sci 101(1):402–407
120. Ting HM, Wang B, Ryden AM, Woittiez L, van Herpen T, Verstappen FW, Ruyter-Spira C, Beekwilder J, Bouwmeester HJ, van der Krol A (2013) The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. New Phytol 199(2):352–366
121. Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD (2009) High-level production of amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. PLoS ONE 4(2):e4489
122. van Beilen JB, Holtackers R, Lüscher D, Bauer U, Witholt B, Duetz WA (2005) Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium* sp. Cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*. Appl Environ Microbiol 71 (4):1737–1744

123. van Herpen TW, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J (2010) *Nicotiana benthamiana* as a production platform for artemisinin precursors. PLoS ONE 5 (12):e14222
124. Wang Q, Hillwig ML, Wu Y, Peters RJ (2012) CYP701A8: a rice ent-kaurene oxidase paralog diverted to more specialized diterpenoid metabolism. Plant Physiol 158(3): 1418–1425
125. Wriessnegger T, Augustin P, Engleder M, Leitner E, Müller M, Kaluzna I, Schürmann M, Mink D, Zellnig G, Schwab H, Pichler H (2014) Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. Metab Eng 24:18–29
126. Wu S, Jiang Z, Kempinski C, Eric Nybo S, Husodo S, Williams R, Chappell J (2012) Engineering triterpene metabolism in tobacco. Planta 236(3):867–877
127. Wurtzel ET, Quinlan R (2013) Cells and methods for producing lutein. WO2013119552 A2, PCT/US2013/024746. Aug 15, 2013
128. Yokoshima S, Tokuyama H, Fukuyama T (2010) Total synthesis of (+)-vinblastine: control of the stereochemistry at C18′. Chem Rec 10(2):101–118
129. Yu F, Okamoto S, Harada H, Yamasaki K, Misawa N, Utsumi R (2011) *Zingiber zerumbet* CYP71BA1 catalyzes the conversion of α-humulene to 8-hydroxy-α-humulene in zerumbone biosynthesis. Cell Mol Life Sci 68(6):1033–1040
130. Yun C-H, Kim K-H, Kim D-H, Jung H-C, Pan J-G (2007) The bacterial P450 BM3: a prototype for a biocatalyst with human P450 activities. Trends Biotechnol 25(7):289–298
131. Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Britton R, Bohlmann J (2012) Bifunctional cis-abienol synthase from Abies balsamea discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. J Biol Chem 287 (15):12121–12131
132. Zerbe P, Hamberger B, Yuen MMS, Chiang A, Sandhu HK, Madilao LL, Nguyen A, Hamberger B, Bach SS, Bohlmann J (2013) Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol 162(2):1073–1091
133. Zhan X, Han LA, Zhang Y-H, Chen D-F, Simonsen HT (2014) Metabolic engineering of the moss *Physcomitrella patens* to produce the sesquiterpenoids patchoulol and α/β-santalene. Front Plant Sci 5:636. doi: 10.3389/fpls.2014.00636 eCollection 2014
134. Zi J et al (2013) Org Biomol Chem 11:7650–7652